

[illegible][illegible]

[illegible]

[illegible][illegible]

[illegible][illegible]

[illegible]

U00001	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	acta1	
--------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--

[illegible][illegible]


```

misc_feature /note="eps-eps3"
clone_lib="KRC1-5"
1..6858
/note="assembly_fragment:01045
fragment_chain:1
clone_end:5P6
vector_side:left"
misc_feature /note="assembly_fragment:00181
6559..11344
/note="assembly_fragment:00181
fragment_chain:1"
11445..15274
/note="assembly_fragment:01118
fragment_chain:1"
15375..18330
/note="assembly_fragment:00832
fragment_chain:1"
18431..38813
/note="assembly_fragment:00599
fragment_chain:1"
38914..42991
/note="assembly_fragment:00760
fragment_chain:1"
43092..46479
/note="assembly_fragment:00784
fragment_chain:1"
46579..61624
/note="assembly_fragment:00618
fragment_chain:2"
61725..70468
/note="assembly_fragment:01180
fragment_chain:2"
70569..78278
/note="assembly_fragment:00949
fragment_chain:2"
78379..80926
/note="assembly_fragment:00888
fragment_chain:3"
81027..89376
/note="assembly_fragment:00772
fragment_chain:3"
89477..92696
/note="assembly_fragment:00426
fragment_chain:3"
92707..96333
/note="assembly_fragment:00451
fragment_chain:4"
96434..114593
/note="assembly_fragment:00496
fragment_chain:4"
114694..133054
/note="assembly_fragment:00882
clone_end:17
vector_side:right"
BASE COUNT 39396 a 27034 c 27151 g 3061 t 1502 others
ORIGIN
Q007 Match 64.0% Score 4437 DB 2 Length 133054
Host Local Similarity 100.0% Pred. No. 0
Matches 4437 Conservative 0 Mismatches 0 Indels 0 Gaps 0
QY 2178 caatctcgaaggaagctccagatggtgaagaaacctgagagccataactctcag 2537
|||||
Db 4471 CAGTCTCTGAGGAGGACCTTCGATGTAAGCAAAATCTTCAGGCAATGCTGAG 33412
QY 2538 ctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2597
|||||
Db 4411 CTCAGCTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33352
QY 2568 gtccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2657
|||||
Db 4451 GTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 33292

```

```

QY 2568 tgaagccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2717
|||||
Db 33291 TGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 34282
QY 2718 ggaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2777
|||||
Db 33231 GAAATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33172
QY 2778 tgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2847
|||||
Db 33171 TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33112
QY 2838 tccaaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2897
|||||
Db 33111 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33052
QY 2898 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2957
|||||
Db 33051 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32992
QY 2958 gtctctctctctctctctctctctctctctctctctctctctctctctct 3017
|||||
Db 32991 GTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 30002
QY 3018 tctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3077
|||||
Db 32931 TCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32872
QY 3078 gtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3137
|||||
Db 32871 GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32812
QY 3138 gtctctctctctctctctctctctctctctctctctctctctctctctct 3197
|||||
Db 32811 GTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 32752
QY 3198 tgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3257
|||||
Db 32751 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32692
QY 3258 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3317
|||||
Db 32691 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32632
QY 3318 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3377
|||||
Db 32631 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32572
QY 3378 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3437
|||||
Db 32571 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32452
QY 3438 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3497
|||||
Db 32511 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32452
QY 3498 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3557
|||||
Db 32451 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32452
QY 3558 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3617
|||||
Db 32391 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32452
QY 3618 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3677
|||||
Db 32331 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32272
QY 3678 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3737
|||||
Db 32271 TGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32212
QY 3738 ttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3797
|||||

```

[illegible][illegible]

as compressions and repeats. All regions were covered by at least one plasmid subclone or more than one M13 subclone and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, SwissProt, TrEMBL, WPI, WormPeP. Information on the WormPeP database can be found at

<http://www.scripps.edu/ncic/ncic10705/ncicmap10705.htm>
was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

<http://www.scripps.edu/ncic/ncic10705/ncicmap10705.htm>
or prior de Jong. For further details see

<http://www.scripps.edu/ncic/ncic10705/ncicmap10705.htm>

IMPORTANT: This sequence is not the entire insert of clone R011-555124. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone R011-555124 is at 1 in this sequence. The true left end of clone R04-584N17 is at 66109 in this sequence. The true right end of clone R011-9801 is at 55714 in this sequence. Exact coordinates are given.

FEATURES

Source

1. 66208

Organization "Human scaffolds"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

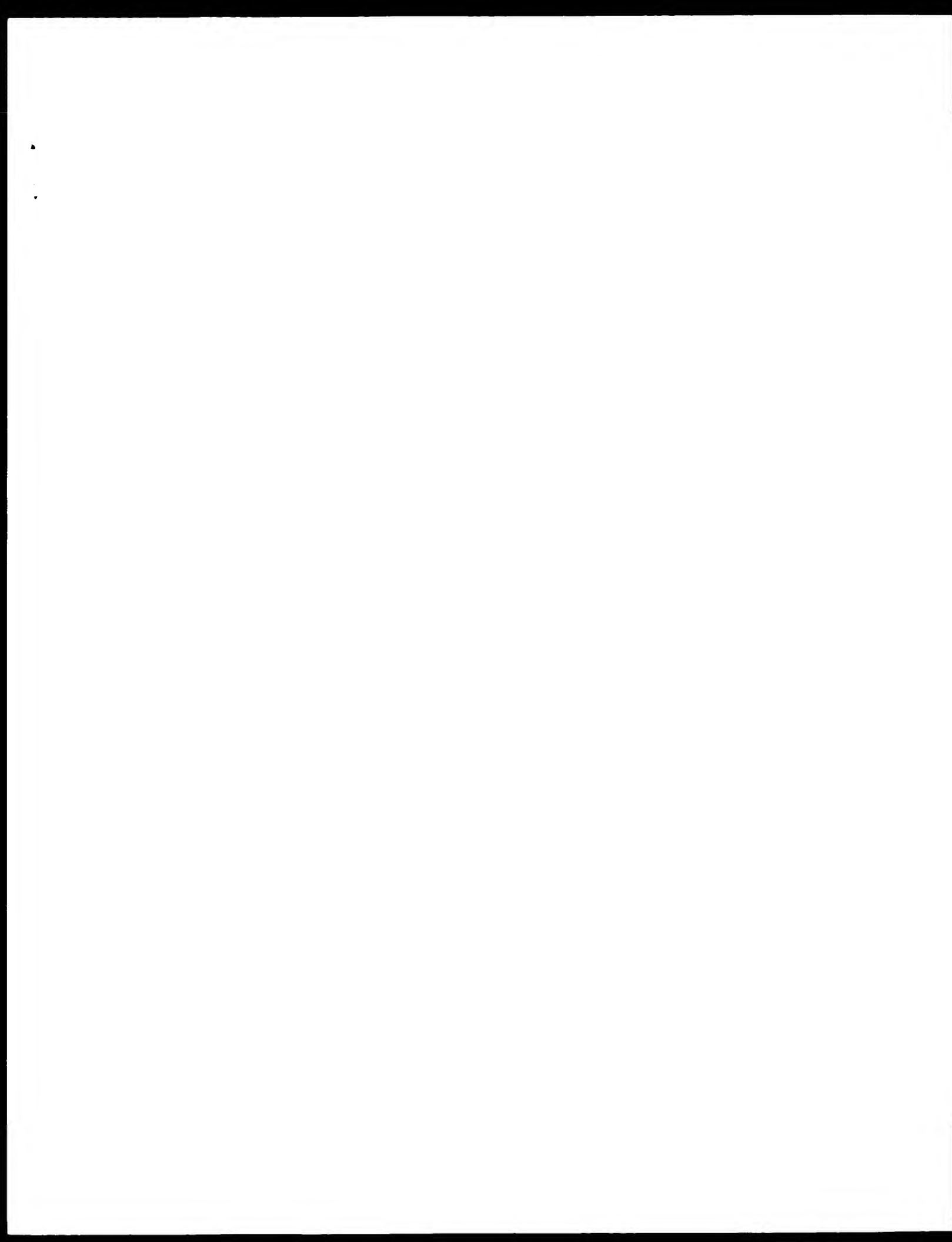
74131..74131..6006"

74131..74131..6006"

74131..74131..6006"

15941	14040:	gap of	unknown	length	
15941	19430:	contig of	3410 bp	in length	
15941	19440:	gap of	unknown	length	
15941	22979:	contig of	5539 bp	in length	
22980	24729:	gap of	unknown	length	
24980	26829:	contig of	3500 bp	in length	
26830	26814:	contig of	unknown	length	
26830	26814:	contig of	1785 bp	in length	
28615	28714:	gap of	unknown	length	
28715	31306:	contig of	3192 bp	in length	
31907	32005:	gap of	unknown	length	
42907	46161:	contig of	4155 bp	in length	
46162	46261:	gap of	unknown	length	
38262	40291:	contig of	4030 bp	in length	
40292	40391:	gap of	unknown	length	
40392	42225:	contig of	1834 bp	in length	
42225	44725:	gap of	unknown	length	
44725	44749:	contig of	4444 bp	in length	
44836	44870:	gap of	unknown	length	
44870	44995:	contig of	4526 bp	in length	
44996	50096:	gap of	unknown	length	
50096	52012:	contig of	2717 bp	in length	
52013	52912:	gap of	unknown	length	
52913	56183:	contig of	3405 bp	in length	
56219	56418:	gap of	unknown	length	
56419	58446:	contig of	2528 bp	in length	
58847	58446:	gap of	unknown	length	
58947	61367:	contig of	2521 bp	in length	
61468	61567:	gap of	unknown	length	
61568	64507:	contig of	2440 bp	in length	
64508	64607:	gap of	unknown	length	
64608	66156:	contig of	1549 bp	in length	
66157	66257:	gap of	unknown	length	
66257	69411:	contig of	4055 bp	in length	
69312	69411:	gap of	unknown	length	
69412	71041:	contig of	2520 bp	in length	
71032	72031:	gap of	unknown	length	
72032	74158:	contig of	2127 bp	in length	
74159	74258:	gap of	unknown	length	
74259	74773:	contig of	2515 bp	in length	
74774	76673:	gap of	unknown	length	
76874	78504:	contig of	1541 bp	in length	
78505	78604:	gap of	unknown	length	
78605	80175:	contig of	1571 bp	in length	
80176	80275:	gap of	unknown	length	
80276	83614:	contig of	335 bp	in length	
83615	83714:	gap of	unknown	length	
83715	85446:	contig of	1632 bp	in length	
85447	85446:	gap of	unknown	length	
85447	86464:	contig of	2139 bp	in length	
86865	86464:	gap of	unknown	length	
86865	87006:	contig of	2402 bp	in length	
89907	90066:	gap of	unknown	length	
90067	92234:	contig of	2719 bp	in length	
92525	92224:	gap of	unknown	length	
92625	94931:	contig of	2457 bp	in length	
94992	95091:	gap of	unknown	length	
95092	97239:	contig of	2148 bp	in length	
97240	97139:	gap of	unknown	length	
97140	99883:	contig of	2744 bp	in length	
99884	99983:	gap of	unknown	length	
99984	102170:	contig of	2187 bp	in length	
102171	102270:	gap of	unknown	length	
102271	103958:	contig of	1608 bp	in length	
104059	104058:	gap of	unknown	length	
104059	105411:	contig of	1953 bp	in length	
105912	106011:	gap of	unknown	length	
106012	107089:	contig of	1928 bp	in length	
107990	108089:	gap of	unknown	length	
108090	109341:	contig of	1252 bp	in length	
109342	109341:	gap of	unknown	length	
109342	110065:	contig of	1375 bp	in length	
110707	110065:	gap of	unknown	length	

[illegible]



1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

Query Match      4.08; Score 204.6; DB 4; Length 7720;
Best local Similarity 79.0%; Pred. No. 6, 96-42;
Matches 282; Conserved 69; Indels 6; Gaps 4;
5767 actctgaattcaatttttttaagatgaatctctctatcaccacattaaa 5826

```


Country Match	2.98%	Score 108,81	108.21	Length 7676;
Best Local Similarity	82.98%	Prod. No. 2,1640;		
Matches 272;	Conserved 120	0;	Mismatches 47;	Indels 5;
				Gaps 2

RESEARCH

COUNTRY: USA
 ZFILE: 10406 0949
 OPERATOR BEHAVIOUR FORM:
 METHOD TYPE: Flapby disk
 COMPUTER: IBM pc compatible
 OPERATING SYSTEM: pc dos/ms dos
 SOFTWARE: Patient In Rollbase #1.0, Version #1.00
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PC7/0595/06/743
 FILING DATE:
 CLASSIFICATION:
 ORDER APPLICATION DATA

[illegible]

1
2
3

Journal: from human brain
 MEDLINE: DNA Res. 4 (5), 345-349 (1997)
 FEATURES: 98116662

SOURCE: Location/Qualifiers

1..6833
 /organism "Homo sapiens"
 /db_xref "taxon:9606"
 /chromosome "1"
 /clone "H06094"
 /sex "male"
 /tissue_type "brain"
 /clone_1db "pblucscript1 SR plus"
 1..2538
 /gene "K1A00457"
 1..2538
 /gene "K1A00457"
 /odon_start 1
 /product "K1A00457 protein"
 /protein_id "K1A0457.1"
 /db_xref "GI:3413876"
 /translation "AAQERREELAACRMPGCGQGAFAAGGEGSVHAGSPDCTPPA
 ACERRRLARPGVYMSSTGCGISPAVGTLPREGVSGESHHSERAKQGLD
 SPGLTVSPVSKSAATVTYVSGTSHRGTQIGETRLPRLSPNGGSRMGOEP
 AADSSPTDASWFAACSHKARKVKAASLPASLSNSGSPGCPVPTPGSSA
 FTSSEPTDLSACHPKAPKAPPSAFSPHQSPEWAKAASLDHPHPLIS
 QPSTLATRVSAIDQAARNSRPEDRMHSLPMDGSSSLDPSLACGCGSSGSG
 DAIWSTLATRVSAIDQAARNSRPEDRMHSLPMDGSSSLDPSLACGCGSSGSG
 LEIDFQKTSIHQGLSPGALSTGLAAQVAALPGATGQASDPVPTLPEP
 RLLEPTGQSLSHITRQMLDQKGLQETVALQAMVLEAKVQALPRLPEEO
 QIDMGQDLPVQGLSDQGVSKAQGLTASVQDPRHAPPTTRSLGRTKST
 NLSKELTKVSKERFSTLRKVNQIPQIDALFAKMHATSGNHWTAKDITEL
 KSLTSPHGLGLSNLVLSSKVNKLSVLYNLRREVHQETAVETSVENIM
 KMYELLNKLTSSKPLAKVWALDLA"ELLQSLQGLQAGSLSEVERQMDLEL
 AAPPILPRLSEIKRFTLKESYDLSAEIPEETLSELTLEYELGLHAIHSHDEL
 LQSLRRELQVKEKTLQMLTQIDPAKENCERFAASGNTAGVHAQA"

BASE COUNT 1861 a 1609 c 1638 g 1725 t

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 6833;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gqctaaagactcaccagaga 21
 |||||

Db 1862 GGCCTAAAGACCTCAGCAGCA 1868

RESULT 4
 LOCUS AX151210 7064 bp DNA linear 22-JUN-2001
 DEFINITION Sequence 1 from Patient W00140101.
 ACCESSION AX151210
 VERSION AX151210.1 GI:14533378
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 7064)
 AUTHORS Porteous,D., Miller,K. and Blackwood,D.
 TITLE A gene, disrupted in schizophrenia
 JOURNAL Patient: W0140101-A 1 07-JUN-2001;
 AKZO Nobel N.V. (NL) ; MEDICAL RESEARCH COUNCIL (GB) ; UNIVERSITY
 OF EDINBURGH (GB)
 FEATURES Location/Qualifiers
 SOURCE 1..7064
 /organism "Homo sapiens"
 /db_xref "taxon:9606"
 BASE COUNT 1930 a 1644 c 1690 g 1799 t
 ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 7064;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gqctaaagactcaccagaga 21
 |||||

Journal: from human brain
 MEDLINE: DNA Res. 4 (5), 345-349 (1997)
 FEATURES: 98116662

SOURCE: Location/Qualifiers

1..6930
 /organism "Homo sapiens"
 /db_xref "taxon:9606"
 /chromosome "1"
 /map "14q2.1"
 1..6930
 /gene "DISC1"
 54..2614
 /gene "DISC1"
 /odon_start 1
 /product "disrupted in Schizophrenia 1 protein"
 /protein_id "AAF73889.1"
 /db_xref "GI:8163869"
 /translation "MGGSPQCAPPAAGGEGSVHAGSPDCTPPAACERRRLARPG;
 YMSSTGCGISPAVGTLPREGVSGESHHSERAKQGLDSPGLTVSPVSKSAATVTYVSGTSHRGTQIGETRLPRLSPNGGSRMGOEP
 AADSSPTDASWFAACSHKARKVKAASLPASLSNSGSPGCPVPTPGSSA
 FTSSEPTDLSACHPKAPKAPPSAFSPHQSPEWAKAASLDHPHPLIS
 QPSTLATRVSAIDQAARNSRPEDRMHSLPMDGSSSLDPSLACGCGSSGSG
 DAIWSTLATRVSAIDQAARNSRPEDRMHSLPMDGSSSLDPSLACGCGSSGSG
 LEIDFQKTSIHQGLSPGALSTGLAAQVAALPGATGQASDPVPTLPEP
 RLLEPTGQSLSHITRQMLDQKGLQETVALQAMVLEAKVQALPRLPEEO
 QIDMGQDLPVQGLSDQGVSKAQGLTASVQDPRHAPPTTRSLGRTKST
 NLSKELTKVSKERFSTLRKVNQIPQIDALFAKMHATSGNHWTAKDITEL
 KSLTSPHGLGLSNLVLSSKVNKLSVLYNLRREVHQETAVETSVENIM
 KMYELLNKLTSSKPLAKVWALDLA"ELLQSLQGLQAGSLSEVERQMDLEL
 AAPPILPRLSEIKRFTLKESYDLSAEIPEETLSELTLEYELGLHAIHSHDEL
 LQSLRRELQVKEKTLQMLTQIDPAKENCERFAASGNTAGVHAQA"

BASE COUNT 1901 a 1625 c 1664 g 1740 t

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 6930;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gqctaaagactcaccagaga 21
 |||||

Db 1862 GGCCTAAAGACCTCAGCAGCA 1868

RESULT 4
 LOCUS AX151210 7064 bp DNA linear 22-JUN-2001
 DEFINITION Sequence 1 from Patient W00140101.
 ACCESSION AX151210
 VERSION AX151210.1 GI:14533378
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 7064)
 AUTHORS Porteous,D., Miller,K. and Blackwood,D.
 TITLE A gene, disrupted in schizophrenia
 JOURNAL Patient: W0140101-A 1 07-JUN-2001;
 AKZO Nobel N.V. (NL) ; MEDICAL RESEARCH COUNCIL (GB) ; UNIVERSITY
 OF EDINBURGH (GB)
 FEATURES Location/Qualifiers
 SOURCE 1..7064
 /organism "Homo sapiens"
 /db_xref "taxon:9606"
 BASE COUNT 1930 a 1644 c 1690 g 1799 t
 ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 7064;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gqctaaagactcaccagaga 21
 |||||

Lib 1862 OCTAATACCTCAGCGAGGA 1882

RESIST 4
A422983

DEFINITION homo sapiens DISC1 protein (DISC1) gene, partial cds and DISC2 gene, partial sequence.

VERSION: 2.2.2983.1
KEYWORDS: GL: 8163835

[illegible]

REFERENCES	1 (pages 1 to 729)
AUTHORS	Miller, J. K., Wilson, Anna, J. C., Anderson, S., Christie, D.,

Blackwood, D. H., and Porteous, D. J.
Disruption of two novel genes by a translocation co-segregating
with a predisposing mutation in the *BRCA1* gene.

JOURNAL. Hum. Mol. Genet. 9 (9), 1415-1423 (2000)
MELLINE 20275640

AUTHORS MILLAR, K., ANDERSON, S. and CHRISTIE, S.
TITLE Direct Submission

FEATURES

11651 13152 contig of 14302 bp in length

13153 13152: gap of 100 bp

13153 146012: contig of 14360 bp in length

146013 146112: gap of 100 bp

146113 166425: contig of 20213 bp in length

Location/Qualifiers

1. 166425

Organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="RP11-39516"

/clone 11b="RP11-11 Human Male BAC"

1. 1252

misc_feature /note="assembly_fragment"

misc_feature 1353..2538

misc_feature /note="assembly_fragment"

misc_feature 2639..3729

misc_feature /note="assembly_fragment"

misc_feature 3830..5231

misc_feature /note="assembly_fragment"

misc_feature 5332..6893

misc_feature /note="assembly_fragment"

misc_feature 6994..8609

misc_feature /note="assembly_fragment"

misc_feature 8710..9718

misc_feature /note="assembly_fragment"

misc_feature 9819..10867

misc_feature /note="assembly_fragment"

misc_feature 10968..13422

misc_feature /note="assembly_fragment"

misc_feature 13523..13543

misc_feature /note="assembly_fragment"

misc_feature 15644..17160

misc_feature /note="assembly_fragment"

misc_feature clone_end:17

misc_feature vector_side:right

17251..21579

misc_feature /note="assembly_fragment"

misc_feature 21480..24663

misc_feature /note="assembly_fragment"

misc_feature 24764..28664

misc_feature /note="assembly_fragment"

misc_feature 28765..33955

misc_feature /note="assembly_fragment"

misc_feature 34056..37279

misc_feature /note="assembly_fragment"

misc_feature 37380..41678

misc_feature /note="assembly_fragment"

misc_feature 41779..46346

misc_feature /note="assembly_fragment"

misc_feature 46447..51724

misc_feature /note="assembly_fragment"

misc_feature 51825..56720

misc_feature /note="assembly_fragment"

misc_feature clone_end:567

misc_feature vector_side:right

56821..65673

misc_feature /note="assembly_fragment"

misc_feature 65774..74120

misc_feature /note="assembly_fragment"

misc_feature 74221..83143

misc_feature /note="assembly_fragment"

misc_feature 83244..93419

misc_feature /note="assembly_fragment"

misc_feature 93520..104494

misc_feature /note="assembly_fragment"

misc_feature 104595..116560

misc_feature /note="assembly_fragment"

misc_feature 116651..141552

misc_feature /note="assembly_fragment"

misc_feature 131653..146012

misc_feature /note="assembly_fragment"

misc_feature 146113..166326

misc_feature /note="assembly_fragment"

RACE COUNT 4776 1 36075 0 46693 1 2828 others

ORIGIN

Query Match 84.8% Score 17.81 DB 21 Length 166325

Best local similarity 90.5% Fred. No. 1.2e+02

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggcctaaagacccacccagga 21

||||| ||||| |||||

Db 111816 ggcctaaagacccacccagga 111836

RESULT 12

AC095057 186927 bp DNA linear HTS 07-09-2001

LOCUS Homo sapiens chromosome 4 clone RP11-39516, WPERIN: JBAFI SEQUENCE,

DEFINITION 16 unordered pieces.

AC095057

AC095057.2 G1:15982599

KEYWORDS HTS: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.

SOURCE human.

ORGANISM Homo sapiens

Fukarya: Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Pluthera; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 186927)

Waterston, R.H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 186927)

Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (15 SEP 2001) Genome Sequencing Center, Washington University, School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

on Oct 7, 2001 this sequence version replaced 9115624971.

COMMENT

Center: Washington University Genome Sequencing Center

Center code: W00SC

Web site: <http://genome.wustl.edu/qsc/index.shtml>

Contact: submissions@wustl.wustl.edu

Project Information

Center project name: H.NH0395106

Summary Statistics

Sequencing vector: M13; 478

Sequencing vector: plasmid; 538

Chemistry: dye-primer ET; 0% of reads

Chemistry: dye-terminator Big Dye; 100% of reads

Assembly program: Phrap, version 0.990319

Consensus quality: 173605 bases at least Q40

Consensus quality: 177343 bases at least Q30

Consensus quality: 179974 bases at least Q20

NOTE: This is a 'working draft' sequence. It currently consists of 16 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1208: contig of 1208 bp in length

1209 1408: gap of unknown length

1309 2640: contig of 1332 bp in length

2641 2740: gap of unknown length

2741 4172: contig of 1432 bp in length

4173 4272: gap of unknown length

4273 6014: contig of 1742 bp in length

6015 6114: gap of unknown length

6115 7316: contig of 1202 bp in length

7317 7416: gap of unknown length

```

* 7417 8785: contig of 1369 bp in length
* 8786 8886: gap of unknown length
* 8886 10245: contig of 1460 bp in length
* 10246 10345: gap of unknown length
* 10346 11861: contig of 1516 bp in length
* 11862 11961: gap of unknown length
* 11962 13546: contig of 1575 bp in length
* 13547 14546: gap of unknown length
* 14547 15156: contig of 1519 bp in length
* 15157 15256: gap of unknown length
* 15257 16614: contig of 1358 bp in length
* 16615 16714: gap of unknown length
* 16715 19113: contig of 2400 bp in length
* 19114 19213: gap of unknown length
* 19214 20880: contig of 1667 bp in length
* 20881 20980: gap of unknown length
* 20981 184505: contig of 163525 bp in length
* 184506 184605: gap of unknown length
* 184606 185481: contig of 876 bp in length
* 185482 185581: gap of unknown length
* 185582 186927: contig of 1346 bp in length.

```

FEATURES

```

source
  1. 186927
    /organism "Homo sapiens"
    /db_xref "taxon:9606"
    /chromosome "4"
    /clone "RP11-39516"
    1. 1208
    /note "assembly_name:Contig12"
    1309. 2640
    /note "assembly_name:Contig18"
    2741. 4172
    /note "assembly_name:Contig30"
    4273. 6014
    /note "assembly_name:Contig31"
    6115. 7316
    /note "assembly_name:Contig32"
    7417. 8785
    /note "assembly_name:Contig33"
    8886. 10245
    /note "assembly_name:Contig34"
    10446. 11861
    /note "assembly_name:Contig37"
    11962. 13546
    /note "assembly_name:Contig38"
    13617. 15156
    /note "assembly_name:Contig39"
    15256. 16614
    /note "assembly_name:Contig40"
    16714. 19113
    /note "assembly_name:Contig41"
    19214. 20880
    /note "assembly_name:Contig42"
    20981. 20980
    /note "assembly_name:Contig43"
    clone_end:589
    vector_side:right"
    misc_feature
      184606..185481
      /note "assembly_name:Contig29"
      185482..186927
      /note "assembly_name:Contig48"
BASE COUNT 53929 a 41155 c 37968 g 52436 t 1549 others
ORIGIN

```

```

Query Match 84.8%; Score 17.8; DB 2; Length 186927;
Best Local Similarity 90.5%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 112453 GCGTAAAGACTCAGCAAGCA 112473

```

```

RESULT 13
AL645615/c 197864 bp DNA linear HTG 01-FEB 2002
LOCUS
DEFINITION
  Mus musculus chromosome 11 clone RP23 395822, *** SEQUENCING IN
  PROGRESS ***, in unordered pieces.
ACCESSION
  AL645615
  AL645615.8 01:18491384
VERSION
  HTG: HTGS_PHASE1: HTGS_ACTIVE1IN: HTGS_DRAFT: HTGS_FULL1DP.
KEYWORDS
  house mouse.
SOURCE
  Mus musculus.
  Fukuoka, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (sites)
  Ashwell, P.
  Direct Submission
  Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  On Feb 4, 2002 this sequence version replaced 91.18476874.
COMMENT
  ----- Genome Center
  Center: Wellcome Trust Sanger Institute
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: humquerry@sanger.ac.uk
  Project Information
  Center project name: BM395822
  Summary statistics
  Assembly program: XGAP4; version 4.5
  Sequencing vector: plasmid; 108752; 100% of reads
  Chemistry: dye-terminator Big Dye; 100% of reads
  Consensus quality: 196787 bases at least Q40
  Consensus quality: 197243 bases at least Q20
  Insert size: 197564; sum-of-consigs
  quality coverage: 16.49x in Q20 bases; sum-of-consigs quality
  coverage: 16.07x in Q20 bases; agafrosc-fp

```

```

* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

FEATURES

```

source
  1. 197864
    /organism "Mus musculus"
    /db_xref "taxon:10090"
    /chromosome "11"
    /clone "RP23-395822"
    /clone_11b "RP11-21"
    1. 127888
    /note "assembly_name:Fragment:02845"
    fragment_chain:1
    clone_end:17
    vector_side:right"
    misc_feature
      127889..154691
      /note "assembly_name:Fragment:03474"
      fragment_chain:1"
      154792..164428
      /note "assembly_name:Fragment:07197"
      fragment_chain:1"
      164529..197864
      /note "assembly_name:Fragment:01640"
      fragment_chain:1"
BASE COUNT 55174 a 43612 c 43370 g 55388 t 420 others
ORIGIN

```

```

Query Match 84.8%; Score 17.8; DB 2; Length 197864;
Best Local Similarity 90.5%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 112453 GCGTAAAGACTCAGCAAGCA 112473

```


•

Average insert size 1.57 kb
Technologies: Note: This is

ORIGIN

Query Match 9.7%; Score 673.4; DB 10; Length 879;
 Best Local Similarity 91.8%; Prod. No. 2.7e-115;
 Matched 80%; Conserved 0; Mismatches 56; Indels 16; Gaps 8.

5950 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 60
 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 60
 6010 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 60
 61 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 120
 6070 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 120
 121 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 180
 6130 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 180
 181 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 240
 6190 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 240
 241 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 300
 6250 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 300
 401 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 360
 6310 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 360
 361 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 420
 6370 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 420
 421 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 480
 6410 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 480
 481 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 540
 6490 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 540
 541 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 599
 6549 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 599
 600 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 659
 6604 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 659
 660 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 719
 6662 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 719
 720 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 777
 6722 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 777
 778 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 832
 6782 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 832
 833 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 867

RESULT 2
 H1756445 962 bp. mRNA linear EST 25-SEP-2001
 DEFINITION 6040502461 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200545 5',
 mRNA sequence.

ACCESSION H1756445
 VERSION H1756445.1 GI:15747923
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens.
 Locus: M1756445.1
 Mammalia: Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 Reference: NIH-MGC http://mgs.mcl.nih.gov/
 Title: National Institutes of Health, Mammalian Gene Collection (MGC)
 Journal: Published (1999)
 Comment: Contact: Robert Strausberg, Ph.D.
 Email: rstraus@nih.gov
 Tissue: fibroblasts, epithelial cells, endothelial cells, etc.
 Tissue procurement: Life Technologies, Inc.
 cDNA library preparation: Life Technologies, Inc.
 cDNA library arrayed by: The J.M.A.G.E. Consortium (J.M.A.G.E.)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the J.M.A.G.E. Consortium/J.M.A.G.E. at: <http://image.jhu.edu>
 Plate: L1M11502 row: 4 column: 10
 High quality sequence start: 11
 High quality sequence stop: 843.
 Location/Qualifiers
 1..962
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5200545"
 /clone_11b="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-Sport6; Site: 1; Note: Site 2: pCMV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yr. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH-MGC Library."

BASE COUNT 224 a 238 c 220 g 279 t 1 others

ORIGIN

Query Match 8.7%; Score 603.4; DB 10; Length 962;
 Best Local Similarity 88.4%; Prod. No. 3e-102;
 Matches 829; Conserved 0; Mismatches 82; Indels 27; Gaps 15;

5189 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 5248
 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 5248
 2 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 60
 5249 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 60
 61 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 118
 5309 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 118
 119 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 176
 5369 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 176
 177 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 236
 5429 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 236
 237 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 295
 5489 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 295
 296 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 355
 5549 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 355
 355 1 TGTATTTTATACACACAGGCTTACCATGTTGGTACAGAGGCTTCACAACTCCGAC 5608

Sufferer human.
 objectivism human sapiens

REPERKUNDE:
Forkarjot; Morozia; Chordata; Crustacea; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases: 1 to 570)


```

bb  450 CCAATATTATCAATTTGATGTTGTAATAAAGCTAAATAATTAATGATGATTAATGCT 291
QY  6783 tcttgatataqtaattcaatcacttttactttaataatcctcacttaactaaacttga 6842
bb  290 TCTTGATAATGTAATCAATCACTTTACTTTAATAATGCTCACTTAATGTAATCTTGAA 231
QY  6843 ttttgatataat'ttattgcttcaatlaagqllactlltllgtatagaataaagaagctg 6902
bb  230 TTTTGCAATGTAATTTATTGCTTCAATTAAGSTTACTTTTGTATACAAATAATAAGCTG 171
QY  6903 atatccaagaca 6914
bb  170 ATATCGACAGCA 159

```

Search completed: September 20, 2002, 06:59:17
 Job time: 10727 sec

BASE COUNT 74 a 39 c 70 g 48 t
ORIGIN

Query Match 82.0% Score 17.4; DB 9; Length 241;
Post Local Similarity 94.7% Pred. No. 2.6e-02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 ctatagactcaccagaa 21
11111111111111111111
14 87 ctAAAGACTCACCAGAA 69

RESULT 4
LOCUS B0460178 400 bp mRNA linear EST 21-APR-2001
DEFINITION B0460178 Athersys KAGE Library Homo sapiens cDNA, mRNA sequence
ACCESSION B0460178
VERSION B0460178.1 GI:14748684
KEYWORDS EST
SOURCE Human
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 400)

REFERENCE
AUTHORS Harrington,J.J., Sherr,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Lovonthal,C., Thornton,M., Ramchandran,R., Whittington,J.,
Lerner,D., Costanzo,P., McElhinny,K., Beozzo,S., Mays,R., Smith,
E., Veloso,N., Kilka,A., Hess,J., Collier,R., Lou,K., Ollendick,
J., Danzig,J. and Inayat,M.
Cloning of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (7), 440-445 (2001)

TITLE
JOURNAL
MEDLINE 21227151
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave., Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com

FEATURES
Source
Location/Qualifiers
1..400
/organism "Homo sapiens"
/db_xref "taxon:9606"
/clone_lib "Athersys KAGE Library"
/note "See description of genome-wide Protein Expression
Libraries using Random Activation of Gene Expression".
Nature Biotechnology, in press. Note that even though the
cell type indicated is H1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in H1080 under normal circumstances."

BASE COUNT 111 a 130 c 83 g 68 t
ORIGIN

Query Match 81.0% Score 17; DB 10; Length 400;
Post Local Similarity 100.0% Pred. No. 4.6e-02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 aaatagctcaccagaa 21
11111111111111111111
14 211 AAAGACTCACCAGAA 227

RESULT 5
LOCUS B0456289 691 bp DNA linear GSS 12-DEC-2001
DEFINITION B0456289 B052 Brassica oleracea genomic clone B052811, DNA
SEQUENCE
ACCESSION B0456289
VERSION B0456289.1 GI:17642000

KEYWORDS GSS,
SOURCE Brassica oleracea,
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE
AUTHORS Tomu,C.D., Van Aken,S., Ultekar,K.T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT other_GSS: B032841R
Contact: Chris Town

TICK
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3524
Fax: 301-838-0208
Email: ctown@nih.gov
DNA is from a doubled haploid provided by Tom Osborn
Seq primer: TF
Class: sheared ends.

FEATURES
Source
Location/Qualifiers
1..691
/organism "Brassica oleracea"
/strain "T01000DH3"
/db_xref "taxon:2712"
/clone "B052811"
/clone_lib "B052"
/note "Vector: PHOS1; SITE_1: B052; 2.4 kb sheared
genomic DNA inserted into pFast1 using balXI linkers"
BASE COUNT 195 a 141 c 191 g 164 t
ORIGIN

Query Match 81.0% Score 17; DB 12; Length 691;
Post Local Similarity 100.0% Pred. No. 5.4e-02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 gctaaagactcaccaa 18
11111111111111111111
14 433 GCTAAGACTCACCAG 417

RESULT 6
LOCUS B0521743 790 bp DNA linear GSS 14-DEC-2001
DEFINITION B0521743 B052 Brassica oleracea genomic clone B0521743, DNA
SEQUENCE
ACCESSION B0521743
VERSION B0521743.1 GI:17729828
KEYWORDS GSS,
SOURCE Brassica oleracea,
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 790)
Town,C.D., Van Aken,S., Ultekar,K.T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
other_GSS: B052174R
Contact: Chris Town

TICK
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3524
Fax: 301-838-0208
Email: ctown@nih.gov
DNA is from a doubled haploid provided by Tom Osborn
Seq primer: TF
Class: sheared ends.

FEATURES
Source
Location/Qualifiers
1..790
/organism "Brassica oleracea"
/strain "T01000DH3"

Feature: Yoo[musculi]
Seq primer: M1 reverse
Location/Qualifiers
1..515

Query Match 80.0% Score 16.8; DB 10; Length 515
Host Local Similarity 90.0% Prod. No. 6.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 481 GCTCAAGACCTCCAGCAGG 500

RESULT 12
B0440944
LOCUS B0440944 566 bp DNA linear GSS 03-DEC-2001
DEFINITION CH240 9b19, IV Chori 240 Segment 1 Ratius novovirus genome clone
VERSION B0440944
KEYWORDS B0440944.1 GI:17271667
SOURCE GSS
ORGANISM Ratius novovirus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
1 (bases 1 to 566)
Author(s): Shetty, J., Shatman, S., Tschagoy, G., Geor, K., Shwarz, Shygn
A., Gabor, G., E., Gabor, L., Russell, D., Chen, D., Riggs, F., de
Jong, P., and Fraser, C.M.
Rat BAC End Sequences from library CHORI-240 Ecoli segment
Unpublished (1999)
Contact: Shyng Shygn
Department of Eukaryotic Genomics
The Institute for Genome Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 401 866 0200
Fax: 401 866 0208
Email: szhagoy@igf.org
Clones are derived from the rat BAC library CHORI-240
(http://www.chori.org/barpac/240.htm). For BAC library
availability, please contact Peter de Jong (pdejong@igf.org).
Clones may be purchased from BACpac Resources
(http://www.chori.org/barpac/or cting_informal_dejong.htm). BAC end
panels: http://www.chori.org/240/bac_ends/rat_end_informal.html
Panel 9, row: F, column: 19
Seq primer: 77
Class: BAC ends
Location/Qualifiers
1..566

FEATURES
SOURCE
1..566
/organism "Rattus novovirus"
/strain "HN/SSNHSd/M/W"
/db_xref "taxon:10116"
/clone "CH240 9b19"
/clone_lib "CHORI-240 Segment 1"
/sex "Female"
/cell_type "Bacul"
/note "Vector: pFACMAC2.1; Site 1: Ecoli; Site 2: Ecoli";

CHORI-240 Rat (HN/SSNHSd/M/W) BAC Library produced by
Peter de Jong
Feature: Yoo[musculi]
Seq primer: M1 reverse
Location/Qualifiers
1..515

Query Match 80.0% Score 16.8; DB 12; Length 566
Host Local Similarity 90.0% Prod. No. 6.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 494 GCTCAAGACCTCCAGCAGG 513

RESULT 13
B1650806/c
LOCUS B1650806 593 bp mRNA linear EST 12-SEP-2001
DEFINITION B01297 99591 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5446175.57,
mRNA sequence.
VERSION B1650806
KEYWORDS B1650806.1 GI:15565042
SOURCE EST
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
1 (bases 1 to 593)
Author(s): NIH-MAC http://mac.cit.nih.gov/
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: chapus@mail.nih.gov
Tissue Procurement: Jathar Hemphausen Ph.D., Chu Xia Tang Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The L.M.A.G.E. Consortium (LMNE)
DNA Sequencing by: Locus Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the L.M.A.G.E. Consortium/LMNE at:
http://image.llnl.gov
Plate: LAM1858 Row: m Column: 24
High quality sequence start: 3
High quality sequence stop: 593.
Location/Qualifiers
1..593

FEATURES
SOURCE
1..593
/organism "Mus musculus"
/db_xref "taxon:10090"
/clone "IMAGE:5446175"
/clone_lib "NIH_CGAP_Mam3"
/tissue_type "tumor, gross tissue"
/lab_host "DH10B"
/note "Organism: mammary; Vector: pCMV-SPORT6; Site 1: Not I;
Site 2: Sal I; Cloned unidirectionally. Primer: oligo dT.
Average insert 2 kb. Library constructed by Life
Technologies, catalog #12017-018. Investigators providing
samples: Jathar Hemphausen/Chu Xia Tang, NIH Reference
for Transgenic model; Xu et al., Nature Genetics 22: 47-43
(1999). Note: this is a NIH_CGAP Library."
Location/Qualifiers
1..566

Query Match 80.0% Score 16.8; DB 10; Length 593
Host Local Similarity 90.0% Prod. No. 6.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 556 GCTCAAGACCTCCAGCAGG 547

RESULT 14
B0608034

ACCESSION VERSION KEYWORDS
 EST.
 H668084.1 GI:10065082
 H668084
 REFERENCE
 AUTHORS
 ORGANISM
 African clawed frog.
 Xenopus laevis
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.
 1 (bases 1 to 595)
 Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Mathis, J., Wyllie, T., Underwood, K., Theising, B., Howers, Y., Peterson, A., Gibbons, M., Harvey, N., Pitter, E., Jackson, Y., Mearns, B., Wilcotton, B. and Wilson, R.
 WashU Xenopus EST project, 1999
 Unpublished (1999)
 Other ESTs: d185q12.x1
 Contact: Sandy Clifton, Ph.D.
 WashU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Library constructed by Bruce Blumberg
 Library normalized by Jimwan Song
 DNA Sequenced by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clone distribution information for this library can be found through Research Genomics. Visit their web page at: <http://www.resgen.com/> Please reference the id listed below when ordering this clone. Source: lab clone id - xlaene003p23
 Seq primer: 40bp from Gibco
 High quality sequence stop. 402.
 Local (on/Qualifiers
 1..599
 /organism "Xenopus laevis"
 /db_xref "taxon:8165"
 /clone "IMAGE:374183"
 /size 595
 /map_pos "Top-10 F"
 /note "Xenopus laevis cDNA from non-normalized 7.5 kb library. XlaI-XhoI cut cDNA was then ligated into Uniap-XR (Stratagene) with EcoRI at the 5' end and XhoI at the 3' end. SS library phagemids were prepared by mass excision from the original library and normalized by hybridization to biotinylated driver (prepared from the same library by PCR) to Cot-omega of 11. After removal of hybrids and excess driver by streptavidin sepharose chromatography, the ss-phagemids were made double stranded and electroporated into Top-10 F⁺ original library construction by Bruce Blumberg (Blumberg et al., 1991 Science 253, 194-196; Hawley et al., 1995, Genes Dev. 9, 2928-2945). Note: This is a Xenopus Gene Definition (XGD) library."
 BASE 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824

AL140359/c	AL140359	677 bp	mean	linear	EST 18-SEP-1997
GENOS	7039466	at	Scarus	11myos 2NM1 Mus musculus cDNA clone IMAGE:577618	
DEFINITION	5' similar to db:X05052 Murine mRNA with homology to yeast 1,29				
	thiosomol protein (pM05E), mRNA sequence.				
ACCESSION	AL140359				
VERSION	AL140359.1	61	1702564		
KEYWORDS	EST.				
SOURCE	house mouse.				
ORGANISM	Mus musculus.				
	Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 677)				
AUTHORS	Marr, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubugny, J.,				
	Gschl, S., Kucaba, T., Lacy, M., Lo, M., Martin, J., Morris, M.,				
	Schulz, P., Schaefer, K., Stephens, M., Tan, F., Thompson, K., Moore, B.,				
	Thompson, B., Wyble, T., Lennon, G., Soares, R., Wilson, R. and				
	Waterston, R.				
TITLE	The Mammal-CHM Mouse EST Project				
JOURNAL	Unpublished (1996)				
COMMENT	Contact: Maria M/Mouse EST project				

WASHU-UMM1 Mouse ES1 Project
Washington University School of Medicine
4444 Forests, Park Parkway, Box 8601, St. Louis, MO 63110
Tel.: 314.286.1800
Fax: 314.286.1810
Email: mouse@watson.wustl.edu
This clone is available royalty-free through UMM1. Contact the
IMM1 Consortium (info@imm1.him.gov) for further information.
MGI: 552566
Seq primer: -28M13 rev2 from Amersham
High quality sequence stops: 475.
Lowell.Dunham@wustl.edu
1-571-6277

```

Zygantism "Mots-motif108"
Zygalin "m57M/bj"
Zyl_xref "Taxon:10090"
Zyloze "IMAGE:577618"
Zyloze_lib "Sources_Thymus_2NMOT"
Zyx "ndic"
Zyssa_type "Thymus"
Zys_slip "4_wcs's"
Zyb_host "DH10B"
Zyote_Vector "pTZ19-Pac (Pharmacia) with a modified
base pairing with a Not I oligo(dI) primer (5'-
TCTGACATCGAAGCAGCAAGCCGGTGTTTTTTTTTTTTTTTTTT-
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia); digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTZ19 vector. RNA
provided by Dr. Bertrand Jordan; library went through two
rounds of normalization, and was constructed by Bontz
Sources and M.Farina Bonaldi."
BASE COUNT      186 a    167 c    141 g    194 t
ORIGIN

```

RESULT: 15

Query Match 77.1% Score 16.2; DB 2; Length 1242;
 Best Local Similarity 85.7% Prod. No. 21;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ggcctaaagacccacacagga 21
 11111111111111111111
 DB 550 ccccaaacacgcacacacga 530

RESULT 2
 US 08 476-062A 52
 ? Sequence 52, Application US/08476062A
 ? Patent No. 5877275
 ? GENERAL INFORMATION:
 ? APPLICANT: Atrium, M. Amin
 ? TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
 ? TITLE OF INVENTION: RESPONSES WITH BETAZ INTEGRINS
 ? NUMBER OF SEQUENCES: 53
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Fish & Richardson P.C.
 ? STREET: 225 Franklin Street
 ? CITY: Boston
 ? STATE: MA
 ? COUNTRY: US
 ? ZIP: 02110-2804
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Diskette
 ? COMPUTER: IBM compatible
 ? OPERATING SYSTEM: Windows95
 ? SOFTWARE: FastSeq for Windows Version 2.0
 ? CURRENT APPLICATION DATA:
 ? FILING DATE: 07 JUN-1995
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: 08/216,081
 ? FILING DATE: 21 MAR-1994
 ? APPLICATION NUMBER: 07/647,830
 ? FILING DATE: 04 JAN-1991
 ? APPLICATION NUMBER: 07/509,842
 ? FILING DATE: 18 JUN-1990
 ? APPLICATION NUMBER: 07/212,573
 ? FILING DATE: 28 JUN-1988
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Friedman, John W.
 ? REGISTRATION NUMBER: 29,046
 ? REFERENCE/CHECKET NUMBER: 0086/068003
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 617/542-5070
 ? TELEFAX: 617/542-8906
 ? TELEX: 200154
 ? INFORMATION FOR SEQ ID NO: 52:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 4704 base pairs
 ? TYPE: nucleic acid
 ? STRANDEDNESS: single
 ? TOPOLOGY: linear
 ? FEATURE TYPE: DNA
 ? US 08 476-062A 52
 Query Match 77.1% Score 16.2; DB 2; Length 4704;
 Best Local Similarity 85.7% Prod. No. 25;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ggcctaaagacccacacagga 21
 11111111111111111111
 DB 1834 ccccaaacacgcacacacga 1834
 RESULT 4
 US 09 134-246-6
 ? Sequence 6, Application US/09134246B
 ? Patent No. 6207377
 ? GENERAL INFORMATION:
 ? APPLICANT: Wayne, Jay
 ? TITLE OF INVENTION: Method for Construction of Thymus E. coli Shuttle
 ? TITLE OF INVENTION: Vectors And Identification of Two Thymus Plasmid
 ? TITLE OF INVENTION: Replication origins
 ? FILE REFERENCE: Thymus Shuttle Vector
 ? CURRENT APPLICATION NUMBER: US/09/134,246B
 ? FILING DATE: 1998-08-14
 ? NUMBER OF SEQ ID NOS: 30
 ? SOFTWARE: Patent In Ver. 2.0
 ? SEQ ID NO 6
 ? TYPE: DNA
 ? LENGTH: 5849
 ? ORGANISM: Thymus sp.
 US-09-134-246-6
 Query Match 77.1% Score 16.2; DB 4; Length 5849;
 Best Local Similarity 85.7% Prod. No. 25;
 Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ggcctaaagacccacacagga 21
 11111111111111111111
 DB 2401 ggcctaaagacccacacagga 2421
 RESULT 4
 US 08 710-249-3
 ? Sequence 3, Application US/08710249
 ? Patent No. 5858777
 ? GENERAL INFORMATION:
 ? APPLICANT: Villiponteau, Bryant
 ? APPLICANT: Feng, Junli
 ? APPLICANT: Andrews, William H.
 ? TITLE OF INVENTION: Methods and Reagents for Regulating
 ? TITLE OF INVENTION: Telomere Length and Telomerase Activity
 ? NUMBER OF SEQUENCES: 26
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Townsend and Townsend and Crew LLP
 ? STREET: Two Embarcadero Center, Eighth Floor
 ? CITY: San Francisco
 ? STATE: California
 ? COUNTRY: USA
 ? ZIP: 94111-3844
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: Patent In Release #1.0, Version #1.0
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/710,249
 ? FILING DATE: 13-SEP-1996
 ? CLASSIFICATION: 536
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 08/583,808
 ? FILING DATE: 05-JAN-1996
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 60/093,492
 ? FILING DATE: 08-SEP-1995
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Storch, John R.
 ? REGISTRATION NUMBER: 32,944
 ? REFERENCE/CHECKET NUMBER: 015389-00122005
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (415) 576-0200
 ? TELEFAX: (415) 576-0300
 ? INFORMATION FOR SEQ ID NO: 3:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 4080 base pairs
 ? TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 79..1380
OTHER INFORMATION: /product= "IPC3"
US-09-770-249-4

Query Match 70.5%; Score 14.8; DB 2; Length 4080;
Best Local Similarity 88.9%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gactaaagactcaccga 18
|||||
DB 3892 GACTAAAGCTACACTA 3809

RESULT 5
US-09-220-157A-3
Sequence 3, Application US/09220157A
Patent No. 6300110
GENERAL INFORMATION:
APPLICANT: Villeporteu, Bryant
APPLICANT: Feng, Junli
APPLICANT: Andrews, William H
TITLE OF INVENTION: Methods and apparatus for localization
TITLE OF INVENTION: Telomere length and telomerase activity
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATOR NUMBER: US-09-220-157A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/710,249
FILING DATE: 13-SEP-1996
APPLICATION NUMBER: US 08/583,808
FILING DATE: 05-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/003,492
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John R.
REGISTRATION NUMBER: 32,944
REFERENCE/POCKET NUMBER: 015389-00122005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 79..1380
OTHER INFORMATION: /product= "IPC3"

US-09-220-157A-3

Query Match 70.5%; Score 14.8; DB 4; Length 4080;
Best Local Similarity 88.9%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gactaaagactcaccga 18
|||||
DB 3892 GACTAAAGCTACACTA 3809

RESULT 6
US-09-446-504-33
Sequence 33, Application US/09446504
Patent No. 6218150
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazuo
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US-09/446,504
CURRENT FILING DATE: 1999-12-23
PRIOR FILING DATE: 1998-06-24
PRIOR FILING DATE: 1998-06-24
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 33
LENGTH: 414
TYPE: DNA
ORGANISM: Pyrococcus furiosus
US-09-446-504-33

Query Match 69.5%; Score 14.6; DB 4; Length 414;
Best Local Similarity 87.8%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gactaaagactcaccga 21
|||||
DB 120 gactaaagactcaccga 140

RESULT 7
US-09-712-266-33
Sequence 33, Application US/09712266
Patent No. 6333158
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazuo
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US-09/712,266
CURRENT FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 09/446,504
PRIOR FILING DATE: 1999-12-23
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496

```

1 PRIOR FILING DATE: 1997-06-26
2 PRIOR APPLICATION NUMBER: JP 9-420692
3 PRIOR FILING DATE: 1997-11-27
4 NUMBER OF SEQ ID NOS: 92
5 SOFTWARE: Patent In Ver. 2.1
6 SEQ ID NO: 43
7 LENGTH: 414
8 TYPE: DNA
9 ORGANISM: Pyrococcus furiosus
10 US 09 712 266 42

```

```

Query Match 69.5%; Score 14.6; DB 4; Length 414;
Best Local Similarity 81.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

CY 1 ggcctaaagactcaccgagaa 21
DB 120 ggcctaaagactcaccgagaa 140

```

```

RESULT 8
US 09 446 504 42
1 Sequence 42, Application US/09/446504
2 Patent No. 6,218,150
3 GENERAL INFORMATION:
4 APPLICANT: DEMORI, Takashi
5 APPLICANT: SATO, Yoshimi
6 APPLICANT: FUJITA, Tomoko
7 APPLICANT: MIYAKE, Kazuo
8 APPLICANT: MUKAI, Hiroyuki
9 APPLICANT: ASADA, Kiyozo
10 TITLE OF INVENTION: DNA POLYMERASE RELATED FACTORS
11 FILE REFERENCE: 1422, 4089CT
12 CURRENT APPLICATION NUMBER: US/09/446,504
13 PRIOR FILING DATE: 1999-12-23
14 PRIOR APPLICATION NUMBER: PCT/JP98/02945
15 PRIOR FILING DATE: 1998-06-24
16 PRIOR APPLICATION NUMBER: JP 9-187496
17 PRIOR FILING DATE: 1997-06-26
18 PRIOR APPLICATION NUMBER: JP 9-420692
19 PRIOR FILING DATE: 1997-11-27
20 NUMBER OF SEQ ID NOS: 92
21 SOFTWARE: Patent In Ver. 2.1
22 SEQ ID NO: 42
23 LENGTH: 1012
24 TYPE: DNA
25 ORGANISM: Pyrococcus furiosus
26 US 09 446 504 42

```

```

Query Match 69.5%; Score 14.6; DB 4; Length 1012;
Best Local Similarity 81.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

CY 1 ggcctaaagactcaccgagaa 21
DB 538 ggcctaaagactcaccgagaa 558

```

```

RESULT 9
US 09 712 266 42
1 Sequence 42, Application US/09/712266
2 Patent No. 6,434,198
3 GENERAL INFORMATION:
4 APPLICANT: DEMORI, Takashi
5 APPLICANT: SATO, Yoshimi
6 APPLICANT: FUJITA, Tomoko
7 APPLICANT: MIYAKE, Kazuo
8 APPLICANT: MUKAI, Hiroyuki
9 APPLICANT: ASADA, Kiyozo
10 APPLICANT: KATO, Kunoshige

```

```

1 TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
2 FILE REFERENCE: 1422, 4089CT
3 CURRENT APPLICATION NUMBER: US/09/712,266
4 PRIOR FILING DATE: 2000-11-15
5 PRIOR APPLICATION NUMBER: US 09/446,504
6 PRIOR FILING DATE: 1999-12-23
7 PRIOR APPLICATION NUMBER: PCT/JP98/02945
8 PRIOR FILING DATE: 1998-06-24
9 PRIOR APPLICATION NUMBER: JP 9-187496
10 PRIOR FILING DATE: 1997-06-26
11 PRIOR APPLICATION NUMBER: JP 9-420692
12 PRIOR FILING DATE: 1997-11-27
13 NUMBER OF SEQ ID NOS: 92
14 SOFTWARE: Patent In Ver. 2.1
15 SEQ ID NO: 32
16 LENGTH: 1012
17 TYPE: DNA
18 ORGANISM: Pyrococcus furiosus
19 US 09 712 266 42

```

```

Query Match 69.5%; Score 14.6; DB 4; Length 1012;
Best Local Similarity 81.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

CY 1 ggcctaaagactcaccgagaa 21
DB 538 ggcctaaagactcaccgagaa 558

```

```

RESULT 10
US 09 418 448 42/1
1 Sequence 42, Application US/09/418448
2 Patent No. 6,210,950
3 GENERAL INFORMATION:
4 APPLICANT: JOHNSON, William G.
5 APPLICANT: STEIGROOS, Edward S.
6 TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
7 FILE REFERENCE: 601-1-057
8 CURRENT APPLICATION NUMBER: US/09/418,448
9 PRIOR FILING DATE: 1999-05-25
10 NUMBER OF SEQ ID NOS: 46
11 SOFTWARE: Patent In Ver. 2.0
12 SEQ ID NO: 42
13 LENGTH: 1256
14 TYPE: DNA
15 ORGANISM: Homo sapiens
16 US 09 418 448 42

```

```

Query Match 69.5%; Score 14.6; DB 4; Length 1256;
Best Local Similarity 81.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

CY 1 ggcctaaagactcaccgagaa 21
DB 686 ggcctaaagactcaccgagaa 666

```

```

RESULT 11
US 09 418 448 45/1
1 Sequence 45, Application US/09/418448
2 Patent No. 6,210,950
3 GENERAL INFORMATION:
4 APPLICANT: JOHNSON, William G.
5 APPLICANT: STEIGROOS, Edward S.
6 TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
7 FILE REFERENCE: 601-1-057
8 CURRENT APPLICATION NUMBER: US/09/418,448
9 PRIOR FILING DATE: 1999-05-25
10 NUMBER OF SEQ ID NOS: 46

```


INDIVIDUAL ISOLATE: CNA2de-11a6 coding sequence
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 262..1767
 US 08 428 122 20

Query Match 69.5% Score 14.6; DB 1; Length 1767;
 Best Local Similarity 81.0%; Prod. No. 1.4e+02;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 US 07 864 475A 4
 US 1152 CCAATTAACACTTCACTCAACA 1172

Search completed: September 20, 2002, 10:34:45
 Job Time: 22219 sec

RESULT 15
 US 07 864 475A 4
 Sequence 4, Application US/07864475A
 Patent No. 5494806
 GENERAL INFORMATION:
 APPLICANT: Sogto, Gino V.
 APPLICANT: Kronenberg, Henry M.
 APPLICANT: Alon Samra, Abdul Radd
 APPLICANT: Juppert, Harald
 APPLICANT: Poats, John T. [Jr.]
 APPLICANT: Schipani, Ernestina
 TITLE OF INVENTION: PAVATHOOLD HORMONE RECEPTOR
 TITLE OF INVENTION: AND DNA ENCODING SAME
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804
 COMPUTER RELEASABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 70Z or 55SX
 OPERATING SYSTEM: MS DOS (Version 5.0)
 SOFTWARE: Wordperfect (Version 5.1)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07864475A
 FILING DATE: 04 06 1992
 CLASSIFICATION: A35
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/681,792
 FILING DATE: 05 04 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Paul L. Clark
 REGISTRATION NUMBER: 40,162
 REFERENCE/PATENT NUMBER: 06786/071002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542 5070
 TELEFAX: (617) 542 8906
 TELEX: 280154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2010
 TYPE: nucleic acid
 SEPARATION: double
 TOPOLOGY: linear
 US 07 864 475A 4

Query Match 69.5% Score 14.6; DB 1; Length 2010;
 Best Local Similarity 81.0%; Prod. No. 1.4e+02;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 US 07 864 475A 4
 US 1152 CCAATTAACACTTCACTCAACA 1172


```

PA (HYSEQ.) HYSEQ INC.
XX
P1 Tang YL, Liu C, Urmanac RL, Asundi V, Zhou P, Xu C, Gao Y, Ma Y,
P2 Zhao CA, Wang D, Wang J, Zhang J, Ren F, Chen K, Wang ZW,
P3 Xue AJ, Yang Y, Wehrman T, Goodrich R.
XX
DE W01: 2001-476284751.
DE P-PSDB: AAM79150.
XX
P1 Nucleic acids encoding polypeptides with cytokine-like activities,
P2 useful in diagnosis and gene therapy.
XX
PS Claim 1: Page 4094-4095; 6221pp; English.
XX
DE The invention relates to polynucleotides (AAK51456-AAK51457) and the
DE encoded polypeptides (AAM78123-AAM78124) that exhibit activity relating to
DE cytokine, cell proliferation or cell differentiation or which may induce
DE production of other cytokines in other cell populations. The
DE polynucleotides and polypeptides are useful in gene therapy, vaccines or
DE peptide therapy. The polypeptides have various cytokine-like activities,
DE e.g. stem cell growth factor activity, hematopoietic regulating
DE activity, tissue growth factor activity, immunomodulatory activity and
DE activity/inhibin activity and may be useful in the diagnosis and/or
DE treatment of cancer, leukemia, nervous system disorders, arthritis and
DE inflammation.
DE Note: Records for SEQ. ID No. 2110 (AAK52581), 2111 (AAK52582) and 3666
DE (AAM80020) are omitted as the relevant pages from the sequence listing
DE were missing at the time of publication.
XX
SQ Sequence 2545 BP; 641 A; 669 C; 750 G; 475 T; 0 other.

Query Match: 100.0%; Score 21; DB 22; Length 2535;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 gactaaagactcaccagaa 21
DB 1848 gactaaagactcaccagaa 1848

RESULT 2
AAK52284
ID AAK52284 standard; cDNA; 2565 BP.
XX
AC AAK52284:
XX
DE 06-Nov-2001 (first entry)
XX
DE Human polynucleotide SEQ. ID No. 828.
XX
DE Human cytokine, cell proliferation, cell differentiation, gene therapy,
DE vaccine, peptide therapy, stem cell growth factor, hematopoietic,
DE tissue growth factor, immunomodulatory, cancer, leukemia,
DE nervous system disorder, arthritis, inflammation; ss.
XX
OS Homo sapiens.
XX
FS Key location/Qualifiers
FS CDS 54..2618 /cds
FT "disrupted in schizophrenia 1 (DIS1) protein"
FE Zproduct
XX
P1 W0200140101-A2.
XX
P2 07-JUN-2001.
XX
P3 28-NOV-2000: 2000W0-EP11915.
XX
P4 01-DEC-1999: 99EP-0409667.
XX
P5 (AIKU) AKZO NOBEL NV.
P6 (MEDIC) MEDICAL RES CORP/ILL.
P7 (OVID) UNIV EDINBURGH.
XX

```

```

PA (HYSEQ.) HYSEQ INC.
XX
P1 Tang YL, Liu C, Urmanac RL, Asundi V, Zhou P, Xu C, Gao Y, Ma Y,
P2 Zhao CA, Wang D, Wang J, Zhang J, Ren F, Chen K, Wang ZW,
P3 Xue AJ, Yang Y, Wehrman T, Goodrich R.
XX
DE W01: 2001-476284751.
DE P-PSDB: AAM79150.
XX
P1 Nucleic acids encoding polypeptides with cytokine-like activities,
P2 useful in diagnosis and gene therapy.
XX
PS Claim 1: Page 2762-2763; 6221pp; English.
XX
DE The invention relates to polynucleotides (AAK51456-AAK51457) and the
DE encoded polypeptides (AAM78123-AAM78124) that exhibit activity relating to
DE cytokine, cell proliferation or cell differentiation or which may induce
DE production of other cytokines in other cell populations. The
DE polynucleotides and polypeptides are useful in gene therapy, vaccines or
DE peptide therapy. The polypeptides have various cytokine-like activities,
DE e.g. stem cell growth factor activity, hematopoietic regulating
DE activity, tissue growth factor activity, immunomodulatory activity and
DE activity/inhibin activity and may be useful in the diagnosis and/or
DE treatment of cancer, leukemia, nervous system disorders, arthritis and
DE inflammation.
DE Note: Records for SEQ. ID No. 2110 (AAK52581), 2111 (AAK52582) and 3666
DE (AAM80020) are omitted as the relevant pages from the sequence listing
DE were missing at the time of publication.
XX
SQ Sequence 2565 BP; 655 A; 672 C; 749 G; 489 T; 0 other.

Query Match: 100.0%; Score 21; DB 22; Length 2565;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 gactaaagactcaccagaa 21
DB 1809 gactaaagactcaccagaa 1829

RESULT 4
AAH24651
ID AAH24651 standard; cDNA; 7063 BP.
XX
AC AAH24651:
XX
DE 06-AUG-2001 (first entry)
XX
DE Nucleotide sequence of a disrupted in schizophrenia 1 (DIS1) gene.
XX
DE Disrupted in schizophrenia 1 gene; DIS1 gene; schizophrenia;
DE psychiatric disorder; ss.
XX
OS Homo sapiens.
XX
FS Key location/Qualifiers
FS CDS 54..2618 /cds
FE Zproduct "disrupted in schizophrenia 1 (DIS1) protein"
XX
P1 W0200140101-A2.
XX
P2 07-JUN-2001.
XX
P3 28-NOV-2000: 2000W0-EP11915.
XX
P4 01-DEC-1999: 99EP-0409667.
XX
P5 (AIKU) AKZO NOBEL NV.
P6 (MEDIC) MEDICAL RES CORP/ILL.
P7 (OVID) UNIV EDINBURGH.
XX

```


XX 10-JAN-2001: 2001MO-US00895.
 XX
 XX 14-JAN-2000: 2000US-0176113.
 XX 21-JAN-2000: 2000US-0177713.
 XX 24-JAN-2000: 2000US-0178832.
 XX 02-FEB-2000: 2000US-0179774.
 XX 03-MAR-2000: 2000US-0186792.
 XX
 XX (INCY) INYTE GENOMICS INC.
 XX
 XX Lal P, Bandman O, Patterson C, Wallis NK, Nguyen EP, Yoo H,
 XX Khan FA, Tang YT, Houghton MP, Lu DM, Yang L, Hartford N,
 XX Ali-Young J, Reddy P.
 XX WPI: 2001-45742/49
 XX P-PSDB: AAE05767.
 XX
 XX Novel isolated human secreted protein useful for diagnosing, preventing
 XX and treating cell proliferative, autoimmune/inflammatory,
 XX cardiovascular, neurological and developmental disorders -
 XX
 XX Claim 5: page 112: 117pt: English.
 XX
 XX The invention relates to human secreted proteins (SECP) and nucleic
 XX acid molecules encoding such proteins. SECP proteins and/or antisense
 XX is useful for treating diseases associated with decreased expression
 XX or overexpression of functional SECP. The disorders treated include
 XX cell proliferative disorders such as cancer, arthritis, osteoporosis,
 XX atherosclerosis, cirrhosis, hepatitis, and psoriasis; autoimmune/
 XX inflammatory disorders such as acquired immune deficiency syndrome
 XX (AIDS), allergies, anemia, asthma, autoimmune thyroiditis, celiac
 XX disease, diabetes mellitus, Goodpasture's syndrome, Graves' disease,
 XX multiple sclerosis, myasthenia gravis, pancreatitis, overactive thyroid
 XX syndrome and infectious cardiovascular disorders such as congestive
 XX heart failure, myocardial infarction, ischemic heart disease and
 XX cardiac hypertrophy; neurological disorders such as stroke, Huntington's
 XX disease, Alzheimer's disease, Pick's disease, epilepsy, dementia, and
 XX Parkinson's disease; and developmental disorders such as renal tubular
 XX acidosis and Cushing's syndrome. SECP DNA is useful to create knock-in
 XX humanised animals (pigs) or transgenic animals (mice or rats) to model
 XX human disease for therapeutic or diagnostic purposes, for screening or
 XX genome therapy, to generate hybridisation probes useful in mapping
 XX the naturally occurring genomic sequence, and in molecular biological
 XX techniques. The present sequence is human SECP cDNA.
 XX
 XX Sequence 1156 BP: 191 A: 388 C: 352 G: 225 T: 0 other:
 XX
 XX Query Match 80.0%; Score 16.8; DB 22: Length 1156;
 XX Best Local Similarity 99.0%; Prod. No. 9;
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX 1 ggcctaaagctcaccagcag 20
 XX 11111111111111111111
 XX 548 ccttAAAGAGCTTACCTGAG 529
 XX
 XX RESULT 9
 XX AA161143
 XX ID AA161143 standard: cDNA: 1167 BP.
 XX
 XX AA161143:
 XX 22-OCT-2001 (first entry)
 XX
 XX Human polynucleotide SEQ ID NO 5132.
 XX
 XX Human neurotrophic immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
 XX leukodystrophy; SS

XX Homo sapiens.
 XX
 XX W020015312-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 26-DEC-2000: 2000MO-US34263.
 XX
 XX 21-JAN-2000: 2000US-0488725.
 XX 25-APR-2000: 2000US-0552317.
 XX 06-JUL-2000: 2000US-0598632.
 XX 19-JUL-2000: 2000US-0620112.
 XX 03-AUG-2000: 2000US-0653450.
 XX 14-SEP-2000: 2000US-0662191.
 XX 19-OCT-2000: 2000US-0693036.
 XX 29-NOV-2000: 2000US-0727344.
 XX
 XX (HYSEQ) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang H,
 XX Wang J, Wang Z, Weinman T, Xu C, Xue AG, Yang Y, Zhang J,
 XX Zhao QA, Zhou P, Goodrich F, Dermanac RT;
 XX WPI: 2001-44745/47.
 XX P-PSDB: AAM11987.
 XX
 XX Novel nucleic acids and polypeptides useful for treating disorders
 XX such as central nervous system injuries -
 XX
 XX Claim 1: SEQ ID NO 5132: 10678pt: English.
 XX
 XX The invention relates to human nucleic acids (AA15799-AA16136) and
 XX the encoded polypeptides (AAM3642-AAW42213) with neurotrophic,
 XX immunosuppressant and cytostatic activity. The polynucleotides are useful
 XX in gene therapy. A composition containing a polypeptide or polynucleotide
 XX of the invention may be used to treat diseases of the peripheral nervous
 XX system, such as peripheral nervous injuries, peripheral neuropathy and
 XX localized neuropathies and central nervous system diseases, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 XX utilisation of the activities such as: immune system suppression,
 XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 XX assays for receptor activity, arthritis and inflammation, leukemias and
 XX C.N.S disorders.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification.
 XX
 XX Sequence 1167 BP: 186 A: 352 C: 361 G: 228 T: 0 other:
 XX
 XX Query Match 80.0%; Score 16.8; DB 22: Length 1167;
 XX Best Local Similarity 99.0%; Prod. No. 94;
 XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX 1 ggcctaaagctcaccagcag 20
 XX 11111111111111111111
 XX 565 ggcctAAAGAGCTTACCTGAG 546
 XX
 XX RESULT 10
 XX AA159357/C
 XX ID AA159357 standard: cDNA: 1169 BP.
 XX
 XX AA159357:
 XX 22-OCT-2001 (first entry)
 XX
 XX Human polynucleotide SEQ ID NO 1560.
 XX
 XX Human neurotrophic immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
 XX leukodystrophy; SS


```

XX  Homo sapiens: 1-3-galactosyltransferase, znssp2, cell-cell interaction.
KW  cancer; autoimmune disease; digestive disorder; gene therapy;
KW  inflammation; chromosome 19q13.2, ss.
XX
XX  Homo sapiens.
XX
XX  Key: Location/Qualifiers
XX  CDS             101..1264
XX                /tag-a
XX                /product: "znssp2"
XX                /transl_except: (Pos:509..511,aa.Xaa)
XX                /note: "Xaa=unknown"
XX  seq_peptide     101..153
XX                /tag-b
XX  mat_peptide     154..1294
XX                /tag-c
XX
XX  W000044486.A2.
XX
XX  15 JUN 2000.
XX
XX  10-DEC-1999: 99WO-0529477.
XX
XX  10-DEC-1998: 98US-0208970.
XX
XX  (Zymo) ZYMOGENETICS INC.
XX
XX  Conklin Inc, Yamamoto C, Jaspers SR, Gao Z,
XX  WPI: 2000 423428/36.
XX  P-PSDR: AAB03619
XX
XX  Novel, znssp2 polypeptides belonging to galactosyltransferase family
XX  useful for treating disorders associated with glycoprotein synthesis.
XX
XX  Disclosures: Page 94-97; 107pp; English.
XX
XX  The present sequence is the coding sequence for human znssp2. This
XX  is a beta-1,3-galactosyltransferase, and is a type II membrane protein
XX  involved in cell-cell interactions. The gene is found on human
XX  chromosome 19q13.2. The gene and protein can be used in the treatment of
XX  cancer, particularly metastases, inflammation, autoimmune diseases such
XX  as diabetes, digestive disorders including Crohn's disease and
XX  ulcerative colitis, and disorders associated with cell migration,
XX  contact inhibition, tissue interactions, neuronal specificity,
XX  fertilisation, embryonic cell adhesion, limb bud morphogenesis,
XX  mesenchyme development, immune response, growth, survival, and
XX  glycoprotein and glycolipid biosynthesis. One possible method of
XX  treatment is the use of gene therapy. The degenerate version of this
XX  sequence is also given (AAA53023).
XX
XX  Sequence 1532 BP: 241 A, 545 C, 459 G, 287 T, 1 other.
XX
XX  Query Match      80.0%; Score 16.8; DB 21; Length 1532;
XX  Best local similarity 90.0%; Pred. No. 86;
XX  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX  1 gcttaagacctcaccagga 20
XX  ||| ||||| ||||| |||
XX  db  gcttaagacctcaccagga 916
XX
XX  RESULT 13
XX  ID AAF61281 standard. RNA. 4800 BP
XX  AC AAF61281;
XX  XX
XX  25-MAY-2001 (first entry)
XX  DE N. magdali bacteriophage phi-CH1 genomic DNA fragment #1.

```

```

XX  Genome, halophilic, polyhydroxybutyrate, inducible expression; ds.
XX  Bacteriophage phi-CH1.
XX  Natrialba magadii.
XX
XX  DE19937719-A1.
XX
XX  22 FEB 2001.
XX
XX  10-AUG-1999: 99DE-1037719.
XX
XX  10-AUG-1999: 99DE-1037719.
XX
XX  (LDB1/) LDB112 W.
XX
XX  Witte A, Baranyi U, Klein R;
XX
XX  WPI: 2001 245330/26.
XX
XX  New nucleic acid from phage phiCh1, used to create vectors for
XX  expressing proteins and polymers in halophilic Archaea.
XX
XX  Claim 1c, Page 16 33, 72pp; German.
XX
XX  This invention describes a novel isolated nucleic acid (1) from the
XX  genome of phage phi-CH1 of Natrialba magadii. The genome of phi-CH1 is a
XX  combination of 4800 and 1000 base pair sequences (s1 and s2), both
XX  fully defined in the specification. The invention also describes a
XX  novel (1) recombinant vector (RV1) containing at least one copy of (1);
XX  (2) recombinant vector (RV2) containing the ori of phi-CH1; (3) cell
XX  transformed with RV1 or RV2; (4) isolated polypeptide (1); encoded by
XX  (1); N. magadii cells free from the prophage of phi-CH1; (5) phage
XX  variants having lytic properties different from those of wild-type
XX  phi-CH1 and having genomic sequences at least 70 % homologous with the
XX  phi-CH1 genome; (6) use of phi-CH1 as gene transfer vector; and (7) use
XX  of halophilic Archaea for production of proteins and other polymers.
XX  Vectors that contain (1) are used to transform halophilic Archaea.
XX  Specifically N. magadii, for production of proteins and other polymers
XX  (e.g. polyhydroxybutyrate). Vectors containing (1) allow controlled,
XX  inducible expression of compounds in Archaea.
XX
XX  Sequence 4800 BP, 10376 A, 14421 C, 15356 G, 8147 T, 0 other;
XX
XX  Query Match      80.0%; Score 16.8; DB 22; Length 4800;
XX  Best local similarity 90.0%; Pred. No. 11e+02;
XX  Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX  2 gcttaagacctcaccagga 21
XX  ||| ||||| ||||| |||
XX  db  gcttaagacctcaccagga 5637
XX
XX  RESULT 14
XX  ID AAS59515 standard. RNA. 66788 BP.
XX  AC AAS59515;
XX  XX
XX  13-FEB-2002 (first entry)
XX  DE Propionibacterium acnes immunogenic protein encoding DNA #10.
XX
XX  SAPHO syndrome; synovitis; acne; pustulosis, hyperostosis, osteomyelitis;
XX  uveitis; endophthalmitis, bone, joint, central nervous system, histiocyte
XX  inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX  dermatological; osteopathic, neutrophilic; ds.
XX  Propionibacterium acnes.
XX
XX  W0200101501-A2.
XX

```

```

PI 01 NOV 2001.
XX
XX 20 APR 2001: 2001W010512865.
XX
XX 21 APR 2001: 2000W051990477.
XX
XX 02 JUN 2000: 2000W052088411.
XX
XX 07 JUL 2000: 2000W052167477.
XX
XX (C9601) CORONA C9601.
XX
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhalla A;
XX L'Herminier J, Zhang Y, Jon S, Carter D;
XX WPI: 2001-616774/71.
XX
XX Proteobacterium genus polypeptides and nucleic acids useful for
XX vaccination against and diagnosing infections, especially useful for
XX treating acne vulgaris
XX
XX Claim 1: SEQ ID NO 10; 1069bp; English.
XX
XX Sequences AAS59506 AAS59804 represent DNA molecules encoding
XX proteobacterium genus immunogenic polypeptides. The proteins and their
XX associated DNA sequences are used in the treatment, prevention and
XX diagnosis of medical conditions caused by P. acnes. The disorders include
XX SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and
XX osteomyelitis), arthritis and endophthalmitis. P. acnes is also involved
XX in infections of bone, joints and the central nervous system, however it
XX is particularly involved in the inflammatory lesions associated with acne
XX vulgaris. A method for detecting the presence or absence of P. acnes in a
XX patient comprises contacting a sample with a binding agent that binds to
XX the proteins of the invention and determining the amount of bound protein
XX in the sample. The polypeptides may be used as antigens in the production
XX of antibodies specific for P. acnes proteins. These antibodies can be
XX used to downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme-linked immunosorbent assay (ELISA).
XX polypeptides shown in A041413.42019 and A0067473.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WPI)
XX at http://wipo.int/pub/publicated\_pat\_sequences.
XX
XX Sequence: 66788 BP; 1554 A; 2169 G; 1947 G; 12530 T; 8 other;
XX
XX Query Match 78.1%; Score 16.4; DB 23; Length 66788;
XX Best Local Similarity 94.4%; Pred. No. 1.8e+02;
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 4 aaagagctcagcagga 21
XX |1111111111111111
XX 19081 aaagagctcagcagga 19098
XX
XX RESULT 15
XX AAC98226
XX ID AAC98226 standard: cDNA; 683 BP.
XX
XX AAC98226;
XX
XX 09 MAR 2001 (first entry)
XX
XX human colon cancer antigen nucleotide sequence SEQ ID NO:26.
XX
XX human; colon cancer; antigen; diagnostic; detection;
XX identification; cytostatic; cardioactive; neuroprotective; vulnerrary;
XX immunomodulatory; muscular; gynaecological; gastrointestinal;
XX nephrotoxic; antitumor; antitumor; antitumor; gene therapy; wound;
XX neural disorder; immune system disorder; muscular disorder;
XX reproductive disorder; gastrointestinal disorder; renal disorder;
XX infectious disease; cardiovascular disorder; ss;
XX

```

```

OS Homo sapiens
XX
XX W0200055451-A1.
XX
XX 21 SEP 2000.
XX
XX 08 MAR 2000: 2000W010505884.
XX
XX 12 MAR 1999: 9905-0124270.
XX
XX (HUMA-) HUMAN GENOME SET 1NC.
XX
XX Rosen CA, Rubin SM;
XX
XX WPI: 2000-587544/55.
XX
XX P PSDB: AAB54469.
XX
XX colon cancer associated gene sequences, referred to as colon cancer
XX antigens, useful for the treatment, prevention, and diagnosis of colon
XX disorders such as colon cancer
XX
XX Claim 1: Page 657: 2104pp; English.
XX
XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,
XX called human colon cancer antigens, given in AAB54244 to AAB54606. The
XX human colon cancer antigens can have cytostatic, cardioactive, muscular;
XX neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
XX vulnerrary, nephrotoxic, antitumor and antitumor polypeptides, and
XX can be used in gene therapy. The colon cancer antigen polypeptides,
XX proteins and antibodies to the proteins are useful for the prevention,
XX treatment and diagnosis of colon disorders, such as colon cancer. The
XX polypeptides may be used in diagnostics and research. The proteins
XX chromosome identification, and as hybridisation probes. The proteins
XX may also be used to prevent diseases such as neural disorders, immu-
XX system disorders, muscular disorders, reproductive disorders,
XX gastrointestinal disorders, wounds, renal disorders, infectious
XX diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
XX AAB54007 represent sequences used in the exemplification of the present
XX invention.
XX
XX Sequence: 683 BP; 143 A; 210 G; 184 G; 136 T; 10 other;
XX
XX Query Match 77.1%; Score 16.2; DB 21; Length 683;
XX Best Local Similarity 85.7%; Pred. No. 1.6e+02;
XX Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 1 qactaaagctcagcagga 21
XX |1111111111111111
XX 275 qactaaagctcagcagga 295
XX
XX Search completed: September 20, 2002, 10:47:58
XX Job Time: 22622 sec
XX

```

Fri Sep 20 16:27:54 2002

us-09-770-107-33.rng

Page 11

•
•
•